

## **Schlagwörter beim Wort genommen**

### **Verbale Erschließung in Bibliothekskatalogen - linguistisch betrachtet**

Gerhard Stumpf

#### **Erschließung und Retrieval mit sprachlichen Mitteln**

Bei der verbalen Inhaltsererschließung (Indexierung), wie sie u.a. Bibliotheken praktizieren, geht es - im Anschluss an den grundlegenden Vorgang der Inhaltsanalyse<sup>1</sup> - um die Erzeugung suchbarer sprachlicher Einheiten (Wörter, Namen, Phrasen), die in Indexen für die Suchabfrage und Anzeige aufbereitet werden und außerdem den bibliographischen Beschreibungen von Dokumenten in einem Katalog hinzugefügt werden.

Mit der Inhalts- oder Sacherschließung soll das Auffinden (retrieval) der bibliographischen Nachweise durch einen Katalogbenutzer ermöglicht werden, der sich "aus der Literatur" über ein mehr oder weniger bestimmtes Thema informieren will. Dass dieses Erschließungsangebot und seine Annahme durch nach Literaturnachweisen suchende Benutzer einen Vermittlungs- und Verständigungsprozess darstellt, dessen Gelingen nicht zuletzt auf linguistischen Kriterien beruht, soll hier in Erinnerung gerufen werden.<sup>2</sup>

Im Vordergrund der Betrachtung stehen die in deutschsprachigen Bibliothekskatalogen - heute i.d.R. Online-Datenbanken - verwendeten Schlagwörter; die allgemeinen Aussagen gelten jedoch ebenso für ähnliche Verfahren, auch in anderen Sprachen.

Unter der Bezeichnung *Indexat* sollen die einzelnen inhaltskennzeichnenden Einheiten (Deskriptoren, Schlagwörter im engeren Sinn) verstanden werden, aber auch die Gesamtheit der ein Thema repräsentierenden (z.B. mehrere verknüpfte) Schlagwörter.

Beispiel:     *AIDS ; Medizinsoziologie*  
                  *Student ; AIDS ; Einstellung*

Schlagwörter repräsentieren Begriffe, in denen sich oft komplexe Inhalte auf zentrale Themen komprimieren und z.T. abstrahieren lassen. Bei den meisten Dokumenten (den nachzuweisenden und zu beschreibenden Publikationen) ist eine solche Aufbereitung bereits von Autoren oder Verlagen durch die Formulierung von Sachtiteln - das sind vielfach Nomen oder Nominalphrasen (z.B. "*Medienidentitäten*", "*Islamische Organisationen in Deutschland*"), auch Eigennamen (allerdings in jeweils unterschiedlichen Sprachen und rein deskriptiv, also ohne normierende Absicht) - teilweise vorweggenommen. Auch durch das Verfassen von Abstracts, kurzen Inhaltsangaben in natürlicher Sprache (i.d.R. Fachsprache), sollen die Inhalte durch die Autoren oder andere Personen so knapp zusammengefasst werden, dass eine kursorische Kenntnisnahme vom Dokumentinhalt möglich wird.

Die Schlagworterschließung ist im Spektrum anderer verbaler Inhaltsererschließung insofern eine Besonderheit, als alle sinntragenden Wörter von Anfang an zu eigenen und oft hervorgehobenen Sucheinstiegen im Katalog bestimmt werden, was für die Wörter in Dokumenttiteln und Abstracts durchaus nicht gilt (unabhängig davon, dass alle Wörter einer Volltextsuche zugänglich sein können). Die einzelne sinntragende Einheit erhält eine Schlüsselfunktion vor allem dadurch, dass sie als Zugriffspunkt bei gleichen Inhalten

mehrerer Dokumente immer die gleiche Form und Aussage haben muss. Wegen dieses großen Gewichts wird sie einer sprachlichen Normierung unterzogen und dieses Erschließungsvokabular in einem Thesaurus (Schlagwort-Normdatei) zusammengefasst.<sup>3</sup>

Untersucht man Form und Funktion der bibliothekarischen Schlagwörter, insbesondere der im deutschsprachigen Raum nach dem Regelwerk RSWK<sup>4</sup> vergebenen, nach den linguistischen Aspekten ihrer Vermittlungsfunktion, so kann man solche in unterschiedlichen Bereichen finden:

- Es gibt den begrifflich-semanticen Aspekt, dem in Regelwerken traditionell große Aufmerksamkeit gewidmet wird und der besonders relevant ist z.B. für die in einem Thesaurus zentrale Frage nach der Synonymie sowie für die Unterscheidung von Homonymen. Eine terminologisch kontrollierte Indexierung setzt eine saubere Begriffsklärung und Disambiguierung voraus. Zur Klärung der Wortbedeutung bedarf es des ständigen Rekurses auf die reale Welt der dargestellten Objekte und Sachverhalte, und zwar auf dem Weg über die abstrakte Welt der Begriffe.
- Vom semantischen lässt sich ein lexikographischer Aspekt unterscheiden, der z.B. alle Fragen der Wortformen umfasst; auf der Ebene des einzelnen Schlagworts ergeben sich hier für die Regeln zentrale Anforderungen, zumal die deutsche Sprache eine große Bandbreite der Präkombination von Lexemen (z.B. Wortbildung) kennt, um Begriffe wiederzugeben. Auch die Möglichkeiten morphologischer Markierungen in Mehrworttermini (z.B. *Feministische Philosophie*) spielen dabei eine Rolle. Die Verwendung von Singular oder Plural, natürlichen und grammatischen Geschlechtsbezeichnungen usw. müssen ebenfalls geregelt sein.
- Indexate besitzen eine syntaktische oder auch textlinguistische Dimension, weil mehrere Schlagwörter zur Beschreibung eines Sachverhalts verknüpft werden können (oder z.T. müssen),<sup>5</sup> wobei zu bedenken ist, dass es sich auf der Ebene von Schlagwortketten - im Gegensatz zu Dokumenttiteln und Abstracts - nur um rudimentäre Formen von Syntax zur Darstellung grammatischer Relationen handelt. Die RSWK sehen keine syntaktische Markierung auf morphologischer Ebene (z.B. wechselnde Flexionsendungen) vor, sondern einheitlich den Nominativ Singular von Substantiven, wenn dieser überhaupt gebräuchlich ist. Ebenso wenig gibt es nach RSWK Rollenoperatoren, also Funktionswörter zur Herstellung einer expliziten und eindeutigen syntaktischen Beziehung in einer Schlagwortkette. Man könnte bei den RSWK von einer parataktischen Struktur sprechen.<sup>6</sup> Umso wichtiger ist, dass aus Schlagwortketten auch ohne explizite Syntax Textbezüge erkennbar werden können. Anfangs legten die RSWK großen Wert auf die Vermeidung des „Pleonasmus“ in der Kette, was mit der 3. Auflage von 1998 mit Recht zurückgenommen wurde, da in Online-Katalogen Konsistenz und Integrität des einzelnen Schlagworts oberste Priorität zukommt. Der sinnvolle Umgang mit dem „Pleonasmus“, dessen bibliothekarische Definition („überflüssige Häufung sinngleicher oder sinnähnlicher Ausdrücke“)<sup>7</sup> von der stilistischen Bedeutung des Terminus durchaus abweicht, war zwar eine entscheidende Voraussetzung für die Anwendung der RSWK in Online-Datenbanken, dürfte aber ein Nebenaspekt sein, wenn Schlagwortketten in ihren Möglichkeiten als sprachliche Einheit zu sehen sind. Indem sie Konnotationen verstärken und Assoziationen freisetzen, haben komplexere Indexate mehr Aussagekraft als die Summe ihrer Teile und eröffnen weitere konzeptuelle und verbale Zugänge zu den Inhalten.<sup>8</sup>

Beispiel: *Italien ; Fürstenkind ; Kleidung ; Geschichte 1600-1700*

*Kinderkleidung <Motiv> ; Malerei ; Geschichte 1600-1700  
Italien ; Hof ; Kinderkleidung ; Geschichte 1600-1700*

Es fragt sich dann auch, ob man Schlagwortketten gar als eigene Textsorte ansehen darf, was für das funktional verwandte Abstract (Inhaltsreferat) ohne Weiteres möglich ist. Fremdindexate wird man allerdings kaum in die Gruppe der "Paratexte" einordnen können, und den Textbegriff überhaupt darauf anzuwenden erfordert dessen weite Definition als Aussageeinheit im kommunikativ-pragmatischen Sinn. "Im Grunde geht es darum, die sachlogische Begriffsabfolge, die der Indexierer geistig präsent hat, zu sichern und niederzulegen."<sup>9</sup> Dann kann jedes Indexat, schon das einzelne Schlagwort, das ja - analog zu Publikationstiteln - ein extremes Kondensat des Dokumentinhalts darstellt, als eine Form von "Metatext" verstanden werden.

- Keineswegs zu vernachlässigen ist der lesetechnische Aspekt, besonders angesichts des der Schlagwortvergabe inhärenten Zwangs zur Beschränkung, die ökonomisch geboten und ergonomisch und in kognitionspsychologischer Hinsicht sinnvoll ist. "Übersichtlichkeit" ist ein Ziel, das nicht nur durch die Anordnung der Texte auf Bildschirmen, sondern auch durch Ansetzungsregeln und am besten durch eine adäquate Abstimmung beider erreicht werden kann. Mit Regelungen, welche Wörter als besonders wesentlich die bevorzugte Stelle am Anfang eines Indexats einnehmen sollen, verfolgt die Schlagwortkatalogisierung das Ziel, habitueller Lesepraxis bzw. einem methodischen Nachschlageverhalten zu entsprechen und die kursorische Lektüre der Inhaltsbeschreibungen, eine Art "flächiger Durchdringung"<sup>10</sup> z.B. in Listen zu erleichtern.

Die Indexierungsarbeit einschließlich der Aufstellung geeigneter Regeln und der Thesauruspflege sollte auf diesen und anderen linguistische Erkenntnissen aufbauen, denn "als Referenzwissenschaft der Inhaltserschließung kommt der Linguistik eine herausragende Rolle zu".<sup>11</sup> Andererseits kann es auch keine erfolversprechende Suchstrategie geben, die nicht auf die sprachliche Form der Erschließung Rücksicht nimmt. Abgesehen von tieferen sprachlichen Einsichten, die weder bei der bibliothekarischen Arbeit noch bei der Literatursuche vorausgesetzt oder als Ziel gesetzt werden können, ist aber eine solide Beherrschung von Rechtschreibung und Grammatik für beide Seiten der beste Weg, um beim Retrieval unnötigen sprachlichen Hürden trivial-handwerklicher Art aus dem Weg zu gehen. Versuche, durch "idiotensichere" OPACs jegliches sprachliche Vorwissen bei Benutzern entbehrlich zu machen, wären zumindest im wissenschaftlichen Bereich der falsche Weg, weil man so leicht hinter den inhärenten Ansprüchen des referierten Wissens zurückbleiben würde.

Die linguistische Relevanz jeder bibliographischen Erschließung liegt schließlich nicht nur in der Verwendung verbaler Mittel, sondern auch darin, dass hierbei mit sprachlichen Mitteln und durch den Rekurs auf sprachliche Kompetenz ein Zweck erreicht werden soll, der letztlich im Auffinden wiederum meist sprachgebundener Werke ("Literatur", "Medien", "Texte", "Dokumente" o.ä. genannt) besteht. Zwischen dem Charakter dieser Zielobjekte und dem für Erschließung und Suche eingesetzten Mittel Sprache besteht ein grundlegender qualitativer Zusammenhang.<sup>12</sup>

### **Wie lernt man Schlagwortdeutsch?**

Die Basis der Schlagwortsprache ist die natürliche Sprache. Nicht-natürlichsprachliche Elemente (z.B. Formeln, nicht verbale Zeichen) können nicht wiedergegeben werden bzw. werden durch sprachliche Ausdrücke ersetzt. Charakteristisch für die Schlagwortsprache ist

jedoch, dass sie diese natürlichsprachige Basis durch Normierungen im Sinn ihrer Zweckbestimmung zu einer Dokumentationssprache umformt, die sich von jeder einzelnen zur menschlichen Kommunikation gesprochenen oder geschriebenen Sprachform unterscheidet. Nicht selten wurden die dadurch entstehenden Indexate als sprachlich unschön und ungenau kritisiert, obwohl Verformungen wie die in der Lexikographie übliche Inversion (eine umgekehrte Wortfolge, um den zentralen Begriff nach vorne zu stellen, z.B. *Kunstfehler, ärztlicher*) aus den Schlagwortregeln für Online-Kataloge längst verbannt sind. Die Schlagwortsprache soll ja von einer möglichst großen Zahl von Katalogbenutzern verstanden werden und darf nicht verkünstelt werden.

Die Instanzen, zwischen denen das Schlagwort eine Verständigung herstellen will, sind 1. der Indexierer als die erschließende Instanz, der die aus der Inhaltsanalyse eines Dokuments gewonnenen Themen und Begriffe mit Bezeichnungen in Form der Indexate (Indexierungssprache) beschreibt und 2. der in einem Katalog Suchende (Benutzer, Leser) mit seiner "Suchsprache". Nun muss die Verständigungsbasis auf fast minimalistische Surrogate reduziert werden, und zwar auf beiden Seiten. Schlagwort-Indexate sollen ein möglichst großes deskriptives Potenzial besitzen, und Benutzer müssen ihre zum Teil recht komplexen Suchfragen teilweise auf ein einziges Suchwort "eindampfen". Um wenigstens einen Teil der Lücke schließen zu können, die zwischen den Formulierungen in den Indexaten einerseits und in Ad-hoc-Anfragen andererseits klafft, muss die Inhaltserschließung üblicherweise davon ausgehen, dass der Benutzer für die Suche seine natürliche Alltags- oder Fachsprache nicht unverändert verwendet, sondern diese seinerseits an die Gegebenheiten bibliographischer Erschließungssysteme anpasst, indem er seine Suchanfrage in Form einzelner aussagekräftiger Wörter stellt und nicht z.B. in Form ganzer Fragesätze.

Im Gegensatz zu Fachdatenbankrecherchen, bei denen ein Spezialist (Rechercheur) als Vermittler die Fragen des Endbenutzers in eine spezielle Retrievalsprache übersetzt, sind Online-Bibliothekskataloge für die Endbenutzer bestimmt und sehen eine explizite "Übersetzung" von Suchanfragen für die Eingabe nicht mehr vor. In vielen Fällen verfehlen sich beide Sprachäußerungen, Indexat und Suchwort, zunächst, und es bedarf weiterer Verständigungsversuche. Katalogbenutzer versuchen beispielsweise ihre Suchanfrage iterativ - nur sprachlich oder auch inhaltlich - zu modifizieren und damit die in den Indexaten enthaltenen, zunächst verborgenen Bezeichnungen zu treffen.<sup>13</sup>

Beispiel:      *"rechtsanwalt in Rom"*  
                   ⇒ *"rechtsanwalt in der antike"*  
                   ⇒ *"Geschichte des Rechtsanwalts"*  
                   ⇒ *"Rechtsanwälte bei Martial"*  
                   ⇒ *"Rechtsanwälte Rom"*  
                   ⇒ *"römische Rechtsanwälte"*  
                   ⇒ *"Rechtsanwälte"*  
                   ⇒ *"Anwalt in der Antike"*  
                   ⇒ *"Anwalt im Altertum"*  
                   ⇒ *"Anwalt"*

Natürlich ist die formulierte Suchanfrage als situative Sprachäußerung das Ergebnis eines Umformungsprozesses im Kopf des Suchenden. Am Anfang steht ein oft recht unscharfes oder bezüglich der Fragerichtung unentschiedenes oder gespaltenes Erkenntnis- oder Informationsbedürfnis, das in Begriffe und Bezeichnungen umzusetzen ist. Oft ist auch ein Text vorgegeben, etwa wenn einer Schülerin oder Studentin für die thematische Literatursuche von der Lehrerin eine Aufgabe vorformuliert wurde, mit der sie einen

Bibliothekskatalog befragen muss. Auch klar vorgegebene Formulierungen lassen sich nicht immer unverändert als Sucheingabe verwenden. Ja es kann geradezu fatal sein, wenn zu starr an vorgegebenen Bezeichnungen festgehalten und nicht variiert, also das semantische und lexikographische Umfeld zu wenig einbezogen wird.

An dem obigen Beispiel ist zu erkennen, wie Benutzer im Suchvorgang versuchen, durch Modifikation ihrer Formulierungen diese mit der Erschließungsterminologie der zur Deckung zu bringen. Durch „try and error“, aber auch durch Revision der eigenen Frage und durch Auswertung der erzielten Teilergebnisse sind Lerneffekte im Sinn eines besseren Retrieval möglich.

Da auf der Seite der Indexierung die Basis für den Sucherfolg gelegt wird, liegt hier auch die Verantwortung dafür, dass die nachgewiesenen Themen möglichst adäquat verbalisiert werden. Die Breite des Spektrums potenzieller Katalogbenutzer macht es sinnvoll, dass sich die Schlagwortsprache weitgehend an der Allgemeinsprache in ihrer standardsprachlichen Ausprägung orientiert und nur dort auf fachsprachliche Bezeichnungen setzt, wo allgemeinsprachliche Bezeichnungen weniger gebräuchlich sind. Auch wer eine hochspezielle Fachterminologie nicht beherrscht, soll in der Lage sein, mit landläufigen Bezeichnungen zum Ziel zu kommen. In bestimmten Fachdisziplinen, wo wissenschaftliche Nomenklaturen (geschlossene Benennungssysteme wie z.B. die botanische Nomenklatur) üblich sind, hat die Schlagwortsprache diese in geeigneter Form einzubeziehen. Regionaler Sprachgebrauch und schichtenspezifische Sprache können in Schlagwörtern nur beim Fehlen standardsprachlicher Bezeichnungen Vorrang haben, sollten aber bei allgemeinsprachigen Bezeichnungen nicht ausgeschlossen sein.

Alternative Bezeichnungen der eben erwähnten und weiterer Typen werden in Normdateien als Äquivalente behandelt und eine davon als Ansetzungsform, die anderen als Verweisungsformen so erfasst, dass es in Online-Katalogen möglich ist, den Sucherfolg von der alternativen Verwendung der einen oder anderen Form durch den Benutzer unabhängig zu machen.

Nichtsdestoweniger bleibt die Dominanz der Ansetzungsformen in fast allen Formen der Anzeige, womit die Ansetzung, d.h. die Festlegung der (im Deutschen) gebräuchlichsten Variante eines Schlagworts, ihre Bedeutung behält. Das Prinzip der allgemeinen Gebräuchlichkeit, das im Regelwerk RSWK zentrale Bedeutung hat, bleibt vor dem Hintergrund der heute üblichen Endnutzerrecherche in immer weniger fachspezifischen als fachübergreifenden Datenbanken ohne Alternative; um eine möglichst gute Voraussehbarkeit der Terminologie für eine möglichst große Zahl einzelner Benutzer aus einem sehr heterogenen Publikum zu erreichen. Im Internet frei angebotene Kataloge, die zudem in Metakataloge einbezogen werden, können ihr Publikum nicht mehr auf eine wie auch immer geartete eingegrenzte Community beschränken und müssen mit Zugriffen aus fachfremden und „populären“ Nutzerbereichen rechnen, ja die Erschließung sollte sich aktiv um diese Zielgruppen bemühen.

Eine ständige Anbindung des Schlagwort-Thesaurus an normative linguistische Referenzdaten erfolgt durch die Bezugnahme auf Nachschlagewerke, insbesondere die in den RSWK geforderte Orientierung an den neuesten deutschen Sprachwörterbüchern wie das *Duden Wörterbuch der deutschen Sprache*, Rechtschreib- und Fremdwörter-Duden, also an einem Standardwortschatz. Für die Kennzeichnung von Homonymen im Thesaurus ist das Bezugsvokabular weder das dort bereits vorhandene allein noch der Wortschatz des einzelnen Indexierers, sondern die objektive Gesamtheit der in einer bestimmten Anzahl von

Wörterbüchern nachgewiesenen Bedeutungsvarianten. Diese wird allerdings stark erweitert durch zahlreiche Fachnachschlagewerke und auch außerhalb von Lexika nachweisbare Homonyme, die häufig als Individualnamen auftreten.

Selbstverständlich sind auch für ein Schlagwortvokabular normative Entscheidungen zu treffen, etwa wie man im Einzelfall mit rechtlichen und ethischen Implikationen von Sprache umgeht (Stichwort „political correctness“; so wurde das Schlagwort *Judensau* als Bezeichnung für die historische antisemitische Judendarstellung<sup>14</sup> wieder aus der Schlagwortnormdatei entfernt und durch *Juden* <Motiv> ; *Schwein* <Motiv> umschrieben) oder wie die durch die Rechtschreibreform entstandenen Neuerungen und Varianten aufgenommen werden. Letzlich wird es hier immer darauf hinauslaufen müssen, den Benutzern den Sucheinstieg mit alten und neuen Schreibungen zu ermöglichen, gleich ob eine „alte“ Schreibung nach der Reform zulässig bleibt oder nicht (z.B. *Alleinerziehende Mutter* neben *Allein erziehende Mutter*). In der Zeit vor der Rechtschreibreform von 1994 hatten sich die Schlagwörter meistens auf die eine vom Rechtschreib-Duden sanktionierte Schreibung beschränkt und die Suche mit Alternativschreibungen kaum berücksichtigt.

Bei deutschsprachigen bibliothekarischen Schlagwörtern nach RSWK in der Schlagwortnormdatei (SWD) handelt es sich zum überwiegenden Teil um Eigennamen. Deshalb spielen gelegentlich onomastische Fragen und bei der Ansetzung von Namen Normierungen in deren Struktur eine große Rolle. Was bei dem in Wörterbüchern nachgewiesenen Vokabular wenig Probleme bereitet, kann hier schwierig sein, nämlich bereits die Gattung eines unbekannten Individualnamens zu erkennen. Hierfür muss eruiert werden, welches reale Individuum überhaupt gemeint ist, welche begriffliche „Entität“ vorliegt. Außerdem werden Schlagwörter in diesem Bereich weithin abhängig von individuellen, z.T. kreativ-willkürlichen Eigenbenennungen, die aus persönlichkeits- oder markenrechtlichen Gründen kaum normierend verändert werden können. Auch dem Kriterium einer indexierungspraktisch verstandenen Gebräuchlichkeit unterliegen diese Namen nur wenig, da hier die Ansetzungsregeln der Formalerschließung (z.B. die Form "*Nachname, Vorname*" bei modernen Namen) übernommen werden. Auch in Form historischer Benennungen und Schreibungen kontrastieren diese Individualnamen innerhalb des Vokabulars z.T. mit einer allgemein- und fachsprachlichen Terminologie, die im Wesentlichen dem aktuellen Sprachgebrauch folgt. Die Schlagwortnormdatei versucht die Eigennamen z.T. nach Gattungen mit Hilfe gebräuchlicher Abstrakta, die als Oberbegriffe hinzugefügt werden, zu ordnen und so der lexikographischen Zerfaserung des Vokabulars entgegenzuwirken.

Im Bereich der Namen, die aus allen Kulturen und Sprachen stammen, sind Regelungen wie die für die Transliteration von Namen aus Sprachen mit nichtlateinischen Schriften unverzichtbar, insbesondere wenn sie in einen Kontext deutschsprachiger Schlagwörter integriert werden sollen. In Bibliotheken ist es schon lange üblich, die Ansetzungsregeln für Namen von Personen, Körperschaften und Werktiteln, die bei der bibliographischen Dokumentbeschreibung (Formalerschließung) anzuwenden sind, auf die Ansetzung der Schlagwörter zu übertragen. Bewusste Abweichungen zwischen beiden Erschließungsbereichen erklären sich daraus, dass bei der Formalerschließung in erster Linie von den in den Dokumenten verwendeten (z.B. originalsprachigen) Namensformen ausgegangen werden muss, bei der Inhaltserschließung dagegen die für deutschsprachige Katalogbenutzer als Sucheinstiege geeignetsten Formen gesucht und vielfach eher deutschsprachig gewählt werden, wenn sie so gebräuchlich sind.

Dies alles soll exemplarisch genannt sein für diverse Anlässe, sich bei der Regelung der Schlagwortvergabe mit sprachlichen Gegebenheiten auseinanderzusetzen, die sich aus der Begrifflichkeit der Inhaltsanalyse ergeben.

### **Das "Glücken von Kommunikationsakten"<sup>15</sup> als Ziel**

Nicht zuletzt müssen die Indexate unter pragmatischen Gesichtspunkten betrachtet werden. Ihre Erstellung ist stark funktional bestimmt. Bei der Literatursuche werden sie im Einzelfall situativ genutzt, zum einen durch die Formulierung von Suchanfragen, auf die auch die beim Suchenden bereits vorhandene Kenntnis von Schlagwörtern Einfluss hat, zum anderen beim Lesen und Verstehen (oder Missverstehen) der gezeigten Indexate durch den Benutzer. Letztlich ist das Gelingen dieses konkreten Kommunikationsvorgangs das Ziel jeder Inhaltserschließung.

Schlagwörter werden per se in einen Verwendungszusammenhang gestellt, der ihren semantischen Gehalt überlagert. (Vom Wortsinn ausgehend ist das „Schlagende“ dieser Wörter der Zweck, nämlich etwas auf- oder nachschlagen zu helfen.) Der Verwendungskontext des Bibliothekskatalogs kann die indirekte Kommunikation zwischen indexierenden und recherchierenden Instanzen auch auf der Ebene des einzelnen Indexats stark beeinflussen. Bestimmte Konventionen und technische Gegebenheiten einzelner Katalogsysteme können die Möglichkeiten der Sprachverwendung einschränken oder erweitern. Insofern wird auch das für Schlagwörter zentrale Kriterium allgemeiner sprachlicher Gebräuchlichkeit um Einflussfaktoren ergänzt, die aus den spezifischen Umständen der Katalogsuche stammen. Generell erhält ein als Sucheinstieg angebotenes Schlagwort dadurch einen Gebrauchszweck zugewiesen, dass es für viele Benutzer weder als präziser Ausdruck einer Suchfrage noch als letztgültige Beschreibung eines Dokumentinhalts in Erscheinung tritt, sondern lediglich als vager Ausgangspunkt oder temporärer Orientierungspunkt im Rahmen einer iterativen Suche. Den Prozesscharakter des Sprachgebrauchs verdeutlicht das obige Beispiel zu den Rechtsanwälten im alten Rom. Aber auch das Gebräuchlichkeits-Kriterium auf Wortebene, wie es den RSWK zu Grunde liegt, lässt sich bereits als pragmatische und "benutzerorientierte" Intention verstehen. Die eigene Sprachverwendung auf das Sprach- und Allgemeinwissen von Benutzern abzustellen und eine gemeinsame Verständigungsbasis zu suchen, wie es bei der Schlagwortvergabe geschehen soll, ist grundsätzlich kontext- und situationsbezogenes Sprachhandeln.

Mindestens so stark wie das Sprachwissen des Autors/Wissenschaftlers und des Indexierers sollte für die Inhaltserschließung das potenzielle mangelhafte oder Nicht-Wissen der suchenden Benutzer maßgeblich sein. Ein Datensatz der SWD muss beides widerspiegeln; er hat die korrekte und maßgebliche Terminologie auszuweisen, zugleich aber auch die ungenauen und falschen, jedoch als Suchwörter vorkommenden Bezeichnungen abzudecken.

Auch die Fähigkeit der Indexierungsterminologie zur Anpassung ist für ihre pragmatische Ausrichtung konstitutiv. Ein Thesaurus muss dem Sprachwandel Rechnung tragen und neue Bezeichnungen dann integrieren, wenn sie sich im Sprachgebrauch z.B. der Medien oder der Fachwissenschaft entsprechend durchsetzen.<sup>16</sup>

Feine Bedeutungsunterschiede zwischen verschiedenen Bezeichnungen kann eine Dokumentationssprache wie die deutschsprachige Schlagwortsprache gemäß RSWK aus praktischen Gründen "einebnen", indem sie im Sinn der sog. Quasisynonymie (RSWK § 315) nur eines der bedeutungsähnlichen Schlagwörter zur Verwendung zulässt und die andere

Bezeichnung wie ein echtes Synonym behandelt. Ziel dieses Verfahrens ist es, die Wahrscheinlichkeit zu erhöhen, dass viele Suchende mit ihrem vermutlich ebenfalls nicht hochdifferenzierten Vokabular in den Indexaten fündig werden, d.h. für ihre Frage relevante "Treffer" erhalten. Der Verlust an Präzision für die meist eher seltenen Suchen, die mit einer größeren semantischen Spezifität rechnen, muss in diesen Fällen in Kauf genommen werden. Dieses Verfahren kann als Beispiel dafür dienen, dass die Regeln für die Schlagwortvergabe nicht ausschließlich einer festgefügt Grammatik oder den semantischen Festlegungen von Wörterbüchern folgen dürfen, sondern offen sein müssen für die spezifischen Anforderungen des Kommunikationszusammenhangs, für den sie bestimmt sind.

## **Suchen und Finden, Lesen, Verstehen und Bewerten**

Triviale Auffassungen des Retrievals in Katalogen gehen davon aus, dass "Treffer" lediglich nach dem exact-match-Prinzip erzielt werden, d.h. die Sucheingabe findet sich zeichengleich in einem Indexat wieder, worauf die betreffende Dokumentbeschreibung (Titelaufnahme) als Treffer ausgegeben wird. Die Bezeichnung "Treffer" ist geradezu charakteristisch für dieses zwar in vielen Online-Katalogen dominierende, aber für die Bewertung guter Erschließung doch nicht allein maßgebliche Modell der Datenbankabfrage.

Bei der Schlagwortsuche lässt sich der Effekt auch und gerade in Online-Umgebungen steigern, wenn man dem Benutzer neben der exact-match-Suche eine teilqualifizierte Suche anbietet, die in irgendeiner Form ein Browsing (Stöbern) ermöglicht.<sup>17</sup> Rudimentäre oder nicht exakt zu Treffern führende Eingaben leiten in eine (in der Regel alphabetisch sortierte) Umgebung (wohingegen bei Sachbegriffen eine semantisch orientierte besser wäre), beispielsweise eine Liste von Schlagwörtern oder Schlagwortketten, was die Chance bietet, ein passendes Schlagwort zu entdecken, das zwar nicht exakt der Sucheingabe, wohl aber dem Suchinteresse entspricht.

Hinzu kommen Effekte der serendipity,<sup>18</sup> womit alle willkommenen Funde gemeint sind, die im Kontext oder in Folge einer beliebigen Suchanfrage unvorhergesehen gemacht werden, also die positiven Nebeneffekte, mit denen erfahrene Benutzer von Online-Katalogen, aber auch konventioneller Register beim Stöbern immer wieder rechnen können. Hierfür ist es wichtig, dass Kataloge materialreiche Anzeigoptionen anbieten, die von sprachlichen oder inhaltlichen Bezüge zu einer ursprünglichen Sucheingabe aus die Verzweigung in neue Suchrichtungen erlauben. Technisch ist hier vor allem an Indexate zu denken, die oder deren Teile als Hyperlink angezeigt sind. Von entscheidender Bedeutung für die Nutzbarkeit solcher Effekte ist, wie syntagmatische und paradigmatische Relationen zwischen Schlagwörtern bei aller Knappheit der Indexate erkennbar gemacht werden.

Eine andere, heute in Online-Katalogen weithin genutzte Möglichkeit, den Sucherfolg unabhängig von exact-match-Ergebnissen zu machen, ist die bereits erwähnte Einbindung des Schlagwortvokabulars als Normdatei, in der sämtliche Synonyme und weitere Bezeichnungen in einem möglichst umfassenden semantischen Netz in den einzelnen Datensätzen verankert sind und die durch einen sinnvollen Indexaufbau ermöglicht, dass eine Vielzahl denkbarer Sucheingaben jeweils auf passende Schlagwörter umgelenkt werden, die wiederum zu Dokumentbeschreibungen zu dem betreffenden Thema führen. Diesem Vorteil einer Schlagwort-Normdatei in Online-Katalogen steht allerdings ein Verlust an Transparenz gegenüber, denn die sprachlichen Verknüpfungen werden meistens im Hintergrund ausgewertet; würden sie immer sichtbar gemacht, könnten die Endergebnisse der Suche nicht so rasch geliefert werden.



In diesem Zusammenhang ist anzumerken, dass gerade bei der üblichen Schlagwortsuche in Online-Katalogen der Unterschied zwischen dem sprachlich komplexeren Schlagwort und dem darin enthaltenen Einzelwort im datentechnischen Sinn (= die Zeichenkette zwischen Spatien) spürbar wird. Bei einer Suche in einem Wortindex fällt potenziell mehr Ballast an, weil das vom Suchenden gemeinte Einzelwort auch in diversen Fügungen mit anderer Bedeutung und in Eigennamen vorkommen kann, was aber einem Suchenden, wenn überhaupt, erst bei der Präsentation inhaltlich unpassender Indexate oder bibliographischer Nachweise klar werden kann.

Beispiel "Straße":	<i>Deutschland / Gefahrgutverordnung Straße</i> <i>Augsburg / Bürgermeister-Fischer-Straße ; Stadtplanung</i> <i>Deutsch ; Roman ; Straße &lt;Motiv&gt;</i> <i>Straße von Gibraltar ; Altkarte</i> <i>Petersen, Jan / Unsere Straße</i> <i>Straße / Bewegungsraum ; Kind ; Sozialisation</i> <i>Galloromanisch ; Wortfeld ; Straße</i>
-----------------------	--

Untersuchungen haben gezeigt, dass die Mehrzahl der Benutzer gar nicht nachvollziehen will, auf welchem Weg ein Online-System ihre Suchanfragen verarbeitet, sondern oft eher in Überschätzung der "Intelligenz" von Retrievalsystemen unterstellt, diese könnten Gedanken lesen, und ungeduldig auf die Ergebnisse wartet, deren Relevanz dann meistens aus dem mental-begrifflichen Fragekomplex heraus eingeschätzt wird, ohne die sprachliche Ebene des Retrievals im Einzelnen zu analysieren. Das erklärt, warum nur selten aus einer wirklichen sprachlichen Analyse von Input und Output Aufschlüsse für das künftige Suchverhalten gezogen werden.

Aus den Grundlagen des Online-Retrieval ergibt sich, dass das gesamte Wortmaterial aus der bibliographischen Erschließung durch Verfahren der linguistischen Datenverarbeitung in quantitativer und qualitativer Hinsicht aufbereitet und auch so der Sucherfolg (hier vor allem hinsichtlich des Recall, d.h. der absoluten Zahl relevanter Treffer) gesteigert werden kann, ergibt sich. Mit Hilfe von Parsern und umfassenden Wörterbüchern im Hintergrund können auch bei bereits normierten Schlagwörtern Flexions- und Pluralendungen erzeugt bzw. eliminiert und zusätzliche Wortformen als Sucheinstiege geschaffen werden. Die sprachliche Form der ursprünglich vergebenen und angezeigten Indexate muss davon nicht tangiert werden, so lange die Anreicherung mit Wortformen sich in nicht sichtbaren Indizes niederschlägt.

Auf der anderen Seite kann auch eine Analyse der Sucheingaben von Benutzern in OPACs zur Verbesserung des Erfolgs der künftigen Suchvorgänge genutzt werden: entweder indem man mit diesen Suchwörtern das Schlagwortvokabular anreichert oder indem man die bisherigen Suchanfragen statistisch-probabilistisch auswertet und als Treffer etwa auch Dokumente auswirft, die früher bereits mit ähnlichen Anfragen gefunden wurden.

## **Ausblick**

Qualitativ hochwertige Indexierung benötigt vielfältige Sprachkompetenz, gerade wenn sie für das Retrieval nicht mehr als durchschnittliche Sprachkompetenz erwarten kann. Moderne Online-Kataloge sollten durch verbesserte Such- und Anzeigetechniken in der Lage sein, die

sprachliche Vermittlungsleistung von Indexaten zu unterstützen. Um diese optimieren zu können, muss sie aber vollständig erkannt sein.

Einerseits rückt in großen und ständig weiter wachsenden bibliographischen Datenbanken mit Schlagwort-Normdaten die semantische und formale Integrität des einzelnen Schlagworts in den Vordergrund: der terminologischen Kontrolle wird um der Konsistenz der Daten willen große Aufmerksamkeit zu widmen sein. Andererseits sollte aber auch der Blick auf größere Gesamtheiten in syntagmatischer (z.B. Schlagwortketten) und paradigmatischer Sicht (z.B. in alphabetischen Zusammenstellungen oder Darstellungen semantischer Relationen wie der Unterbegriffe zu einem Oberbegriff) zu den Grundfunktionalitäten von Benutzeroberflächen für die inhaltliche Suche gehören. Schließlich muss auch die Präsentation formalbibliographischer und inhaltlicher Informationen aus kognitionspsychologischer Sicht stärker ganzheitlich gesehen und optimiert werden, damit sich der intendierte Mehrwert der Erschließung in jedem Anwendungskontext erweisen kann.

<sup>1</sup> Inhaltsanalyse (subject analysis) im Sinne der Analyse des thematischen Inhalts von Dokumenten, vgl. z.B. Fugmann, Robert: Theoretische Grundlagen der Indexierungspraxis. Frankfurt/Main 1992; Nohr, Holger: Inhaltsanalyse. In: NfD 50 (2/1999), S. 69-78

<sup>2</sup> Von den wenigen früheren Beiträgen, die die bibliothekarische Inhaltserschließung explizit unter linguistischem Aspekt thematisieren, seien die Arbeiten von Werner Bies genannt: Bies, Werner: Linguistische Pragmatik - eine vernachlässigte Referenzdisziplin der Inhaltserschließung. In: Kognitive Ansätze zum Ordnen und Darstellen von Wissen. Frankfurt/Main 1992, S.207-216; ders.: Das Kriterium 'Gebräuchlichkeit des Schlagworts' in der Geschichte und Theorie des Schlagwortkatalogs. In: Aufbau und Erschließung begrifflicher Datenbanken. Oldenburg 1995, S.19-46

<sup>3</sup> Klassifikatorische Erschließung hat heute im Kontext von Datenbanken und Online-Recherchen ebenfalls eine starke verbale Komponente, die in Katalogsystemen ähnlich wie Schlagwörter eingesetzt werden kann. Den Dokumenten, die inhaltlich zu einer Klasse gehören, werden exakt die gleichen, diese Klasse beschreibenden Texte (in der Regel nur einzelne Begriffe / Wörter) zugeordnet, die sich als Ganzes oder in ihren Bestandteilen für die Suche aufbereiten lassen. Das Vokabular einer Klassifikation unterscheidet sich von einem Thesaurus in der Regel durch eine stärkere Konzentration bzw. Vergrößerung, d.h. bestimmte inhaltliche Aspekte und Spezifika fallen hier weg; dafür kommen zusammenfassende Bezeichnungen hinzu.

<sup>4</sup> Regeln für den Schlagwortkatalog : RSWK. – 3. überarb. und erw. Aufl. [Loseblattwerk. Grundwerk] Berlin : Dt. Bibliotheksinstitut, 1999. - 1. Erg.lfg. 2000. 2. Erg.lfg. 2002.

<sup>5</sup> Zur Funktion von Schlagwortketten als Mehrwert vgl. Flachmann, Holger: Effiziente Sacherschließung als Kernaufgabe von Bibliotheken : allgemeine Probleme und das Fallbeispiel der Regeln für den Schlagwortkatalog (RSWK). 2003. - <http://archiv.tu-chemnitz.de/pub/2003/0065/index.html> , S.11 ff.

<sup>6</sup> Retti, Gregor: „Schlagwortnormdatei“ und „Regeln für den Schlagwortkatalog“. 1995. – [http://germanistik.uibk.ac.at/germ/docs/swd\\_rswk.pdf](http://germanistik.uibk.ac.at/germ/docs/swd_rswk.pdf) , S.4

<sup>7</sup> Regeln für den Schlagwortkatalog : RSWK. 2.Auflage, Berlin 1991, S.9

<sup>8</sup> Vgl. Fugmann, Robert: Theoretische Grundlagen der Indexierungspraxis. Frankfurt/Main 1992, S.216

<sup>9</sup> Flachmann, a.a.O., S.13

<sup>10</sup> Flachmann, a.a.O., S.7

<sup>11</sup> Bies, Werner: Linguistische Pragmatik (s. Anm.2), S.208

<sup>12</sup> Auch innersprachliche Phänomene können Themen von Publikationen sein, die durch entsprechende Schlagwörter erschlossen werden. Beispielsweise können alle Laute (Phoneme), Morpheme, Lexeme (z.B. auch Wortwurzeln, Phraseolexeme) auch durch je eigene Schlagwörter wiederzugeben sein. Beispiel: *Deutsch* ; *Suffix* ; *keit* ; *Wortbildung*. Im Kontext einer RSWK-Schlagwortkette erhalten solche linguistischen Einheiten durch die jeweilige Sprachbezeichnung und die Angabe einer grammatischen Kategorie bzw. Wortart die hinreichende Verständlichkeit.

<sup>13</sup> Reales Beispiel, Recherchemitschnitt aus dem OPAC der Universitätsbibliothek Augsburg vom 17.1.2005

<sup>14</sup> Vgl. hierzu <http://www.christen-und-juden.de/index.htm?html/judensau.htm>

<sup>15</sup> Schmidt, Siegfried J.: Texttheorie/Pragmalinguistik. In: Lexikon der germanistischen Linguistik, hrsg. von H.P.Althaus, H.Henne u. H.E.Wiegand. Tübingen 1973, S.239

<sup>16</sup> Gute Voraussetzungen hierfür bietet die als Online-Datenbank laufend aktualisierte Schlagwortnormdatei (SWD), während Fachthesauri häufig nur in größeren Abständen überarbeitet werden.

---

<sup>17</sup> "Das Blättern (Browsing) im Index der SW-Ketten stellt einen eigenständigen Suchvorgang dar, der die typischen Recherchestrategien des Online-Katalogs, Freitextsuche und Postkoordination, sinnvoll ergänzt und dabei sogar die Schwächen dieser Methoden eliminiert und die Stärken beider in sich vereint." (Weidemüller, Hans Ulrich: Sinn und Unsinn von Ketten im OPAC. 1995. - [http://www.ub.uni-duesseldorf.de/projekte/milos/vortraege/mil\\_weid](http://www.ub.uni-duesseldorf.de/projekte/milos/vortraege/mil_weid) , 4)

<sup>18</sup> Vgl. Fugmann, Robert: Theoretische Grundlagen der Indexierungspraxis. Frankfurt/M. 1992, S.70